



Empirical Analysis of Operation Iraqi Freedom Combat Mortality Using the Navy-Marine Corps Combat Trauma Registry Expeditionary Medical Encounter Database for Applications to Tactical Medical Logistics Modeling and Simulation

R. Mitchell
J. Parker
M. Galarneau
P. Konoske

Naval Health Research Center

Report No. 09-33

. Approved for public release: distribution is unlimited.

Naval Health Research Center
140 Sylvester Road
San Diego, California 92106

Empirical Analysis of Operation Iraqi Freedom Combat Mortality Using the Navy-Marine Corps Combat Trauma Registry Expeditionary Medical Encounter Database for Applications to Tactical Medical Logistics Modeling and Simulation

Ray Mitchell, PhD, TBE
Joe Parker, TBE
Teledyne Brown Engineering, Inc.
300 Sparkman Drive
Huntsville, AL 35805

Mike Galarneau, MS, NREMT, NHRC
Paula Konoske, PhD, NHRC
Naval Health Research Center
140 Sylvester Rd.
San Diego, CA 92106-3521

Report No. 09-33, supported by the Office of Naval Research, under Work Unit No. 60315. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Navy, Department of Defense, or the U.S. Government. Approved for public release; distribution is unlimited. This research has been conducted in compliance with all applicable federal regulations governing the protection of human subjects. No human subjects were directly involved in this research.

Table of Contents

Figures

Tables

Summary

Problem

Reducing U.S. forces' battlefield mortality during combat operations is a paramount concern during operations planning. Methods to more accurately forecast medical needs and effectively meet those needs are constantly being developed to support service-related modeling and simulation (M&S) efforts under the general research topic of tactical medical logistics. The Naval Health Research Center (NHRC) Tactical Medical Logistics (TML+) simulation tool is gaining respect in the planning community. However, further testing and improvement is needed to confirm or strengthen its capabilities to deliver accurate and useful information to planners. It is also expected that an empirically based mortality model would be applicable to other battlefield and civilian tactical medical logistics M&S efforts.

Objectives

This study focused on two main objectives:

1. Describe a research effort to data mine NHRC's Navy-Marine Corps Combat Trauma Registry Expeditionary Medical Encounter Database (CTR EMED) to identify individual patient conditions and the associated risks of mortality.

2. Determine if a probability model could adequately describe the mortality events in a casualty's medical treatment flow as recorded in the Operation Iraqi Freedom (OIF) CTR EMED database. Confirming or updating the expert medical doctor (MD) panel results (Mitchell et al., 2004) with a statistical analysis of empirical mortality results is an obvious research goal, and this effort was directly related to confirming the efficacy of the mortality modeling approach used in TML+.

Methodology

Described within is the data collection effort geared toward identifying life-threatening injury records in the CTR EMED. Then the data are used to develop a probability distribution for the timing of deaths within a medical treatment facility (MTF). Finally the CTR EMED data are used here to examine the efficacy of confirming and/or supplementing the MD panel results with a time-based mortality analysis. Methods from the biomedical sciences and applied life data analysis literature are used throughout.

Results

Statistical analysis results of OIF mortality, for certain resuscitative-capable MTFs and for high risk of mortality patients, showed a strong graphical agreement with results estimated by the 2003 MD SME panel. This confirms the efficacy of TML+ output. Future activities to extend these results across a wider spectrum of the casualty treatment and evacuation path are also discussed.

Conclusions

The statistical analysis and resulting strong graphical agreement between OIF mortality and estimates made by the 2003 MD SME panel confirms the efficacy of TML+ output. Subject matter expert (SME) input for providing missing timing data is a legitimate and valuable way to strengthen biomedical science data records.

Introduction

Reducing U.S. forces' battlefield mortality during combat operations is a paramount concern during operations planning. The Naval Health Research Center (NHRC) Tactical Medical Logistics (TML+) planning tool is capable of simulating a broad range of stochastic events associated with battlefield casualty disposition from the point of injury (POI) to higher levels of care. One of the tool's most important metrics, and the focus of numerous studies in recent years, is died of wounds (DOW) due to delay in treatment estimate (Konoske, 2008). Typical research questions include how many lives can be saved if a forward resuscitative surgical system (FRSS) is located near the combat area?; what is the best way to integrate transportation and medical treatment facility (MTF) assets to reduce DOW incidents in a deployed network of care?; and how do mass casualty events influence operating room throughput and mortality under various staff and equipment configurations? There are no static answers to these questions; the environment and the scenario for each mission are always unique.

Continuous improvements can be made to medical requirements estimation using modeling and simulation. For example, merging cogent subject matter expert (SME) knowledge with this empirically based modeling capability can perhaps improve mortality risk estimates associated with planning operational courses of action, either for pre-planning, or for crisis action planning during real-time deployments. At NHRC, SME knowledge has not been integrated in this way before.

The objective of this research report is to document NHRC's ongoing effort to improve the statistically based modeling of TML+ and other medical logistics systems analysis models. These models project the effects of treatment delays and available resuscitation (surgical-level resuscitation vs. non-surgical-level resuscitation) on battlefield mortality and are a valuable resource for planners.

Background

In TML+, it is assumed that a casualty's time to death is a random variable that has a probability density function (PDF) with parameters dependent on the casualty's injury extent, existing MTF capability, treatment history, and treatment timing. Employing the statistical model, draws are made from the PDF at various stages of a scenario to simulate mortality events. The Weibull PDF currently used in TML+ (to be described later) is based on expert opinion data obtained from a panel of nine military medical doctors (MDs) convened at NHRC in November 2003 (Aylward, 2004; Galarneau & Mitchell, 2003, 2004; Mitchell, et al., 2004).

Figure 1 shows an example of the results obtained from the MD SME panel. In the figure, a casualty receives a series of medical interventions for a life-threatening (LT) injury:

- by self or buddy aid at the point of injury (labeled "no treatment")
- by a field-level corpsman (first responder, 1stR) after a 10-minute delay
- by the battalion aid station (BAS) MTF after a 30-minute delay
- by a Shock Trauma Platoon (STP) MTF after a 30-minute delay

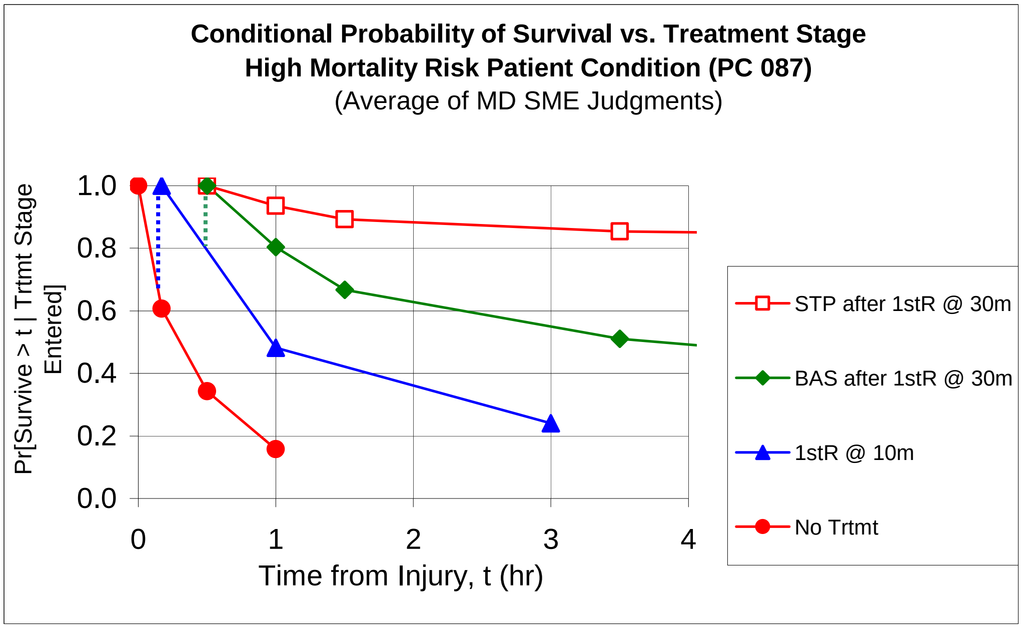


Figure 1. MD SME panel survivability observations for a seriously injured casualty.

The results illustrate the effect of medical treatment level increases as the casualty moves through the MTF system. Curves like this are used in TML+ to simulate time of probable death or survival as a casualty with LT injuries moves through the medical chain of treatment and evacuation. A fit of the Weibull distribution to these same results, estimated by the ad hoc analytical method of matching a few percentile points (Elandt-Johnson & Johnson, 1980), is shown in Figure 2.

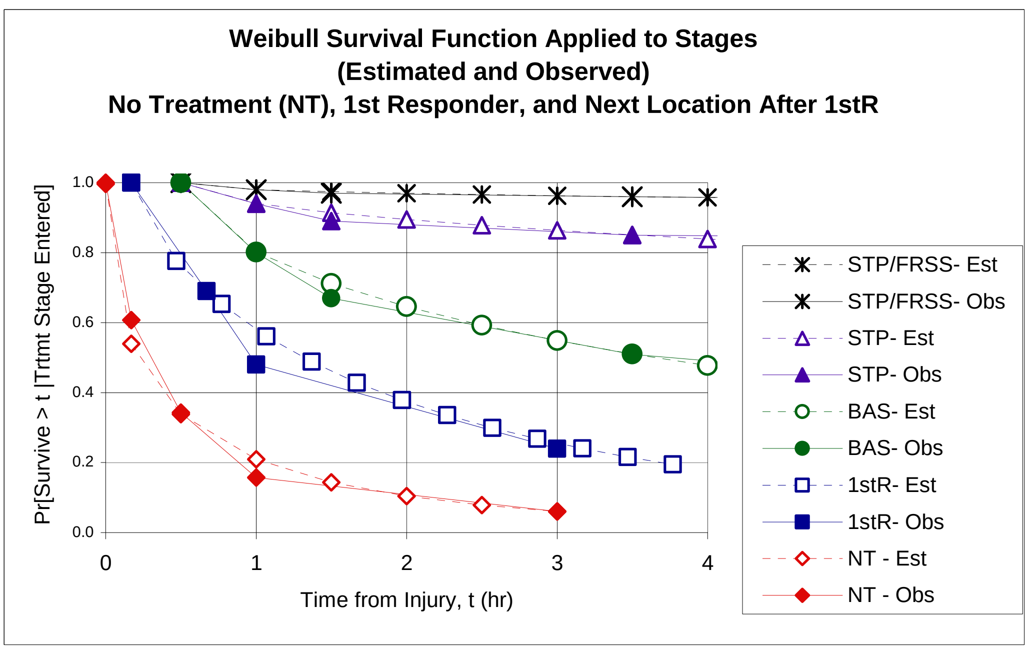


Figure 2. Weibull PDF applied to stages.

Confirming or updating the expert MD panel results with a statistical analysis of empirical mortality results is an obvious research goal. A prerequisite to such analysis is a large data set consisting of LT incidents with complete timing information. Real-world casualty resuscitative data, including data from Operation Iraqi Freedom (OIF), are available in the CTR EMED. These are used here to examine the efficacy of confirming and/or supplementing the MD panel results with a time-based mortality analysis Galarneau, Hancock, Konoske, et al., 2004). The data files used in our analysis cover a period from early 2004 to mid-2007. There were, however, inherent limits encountered in the data that prevent a complete analysis of battlefield mortality as a function of casualty flow paths and delays in treatment in TML+. This is where SME input was very useful in examining suspect data records and/or supplying certain missing data fields.

This report presents the following:

- a description of our data collection efforts to identify LT injury records in the CTR EMED
- statistical analysis results of mortality using techniques from the biomedical sciences and the applied life data analysis literature for certain MTFs
- suggestions for next steps to extend this research and further improve upon the modeling of mortality in TML+ as it relates to treatment delays and existing resuscitative levels of care

Methodology

Data Collection for Mortality Analysis

The CTR EMED is composed of data sets that record medical encounters that occurred to individual casualties from POI through evacuation and beyond (Galarneau et al., 2004). The CTR EMED assists medical planners, systems analysts, and logisticians in planning medical treatment needs for warfighters during deployments. This empirical means of confirming or augmenting medical logistics planning leads to an optimal mix of health care staff, equipment, supplies, and transportation assets being deployed to an area of operations in support of our warfighters. Using the CTR EMED database we will provide a probability distribution for the simulated timing of deaths within an MTF. (This distribution is perhaps dependent on prior treatment paths as currently modeled in TML+.)

Related research to date has concentrated on identifying casualty medical encounters in the CTR EMED that represent LT injuries, which include recorded times for entry and exit in the casualty's medical treatment flow. (By convention, we define LT injuries as those in which a casualty is expected to die within the first hour after injury [the so called "Golden Hour"] if no treatment beyond self/buddy aid is received.) In TML+, LT casualties are identified and designated as having either a high (H), medium (M), or low (L) risk of mortality within the first hour after wounding. If designated (H), the casualty has a probability of dying greater than two thirds; if (M) between one third and two thirds; and if (L) less than one third (Lowe, Hill, & Galarneau, 2008; Mitchell, et al., 2004). Because early CTR EMED records did not contain these designations, we researched the feasibility of implementing data mining means in order to identify LT injuries and assign them a patient condition (PC) code or an (H), (M), or (L) risk designation using an automated algorithm. We attempted to develop an algorithm to facilitate a manual inspection of each record by SMEs to make these designations if we found that a fully automated means was not possible. Later, with an enhanced data set, we planned to test hypotheses with mortality models. We were very interested in confirming the applicability of the Weibull PDF as a reasonable predictor of death times once a casualty had begun to receive medical treatment.

It is useful to note some features of survival analyses typically experienced in biomedical sciences literature that are evident in this effort. Survival analyses require meticulous timing information on each casualty so that a distribution of times of deaths can be found for a statistical analysis. Also, it is common to rely heavily on graphical analysis and nonparametric methods to analyze life lengths (Elandt-Johnson & Johnson, 1980). With this in mind we discuss CTR EMED data and statistics derived from it: wounding characteristics, treatments, evacuations, mortality locations, and timings. Later, we will describe the chronology of our CTR EMED data-mining efforts and the final data set that was obtained. We used these data to further examine the mortality modeling approach and assumptions in TML+. We expect this research will provide sound background material for any future efforts to enhance casualty record interpretation in the CTR EMED database by automated or semiautomated means.

Appendix A shows the CTR EMED form used by field medical providers for recording injury data sustained by warfighters. An individual record captures details regarding care, including death or survival evacuation to the next level of MTF. (These variables are required for the statistical analysis of an individual's risk of death within a population.) The data on these forms are later entered into the CTR EMED database.

Table 1 shows a screen shot of a small subset of the CTR EMED data file. (Note that the timing of injury events and MTF entry/exits, along with disposition data, are suppressed in the table to meet privacy security concerns.) A metric for risk of mortality, such as an Injury Severity Score (ISS), was not included in the data set at the start of this work; it was later added.

Table 1. A Desensitized Portion of the CTR EMED File

Case ID	Medical Treatment Facility (MTF)	Arrival Method	Injury Severity Score	Triage Category	Injury Category	SOAP notes	Disposition	Evac Priority
1111	Battalion Aid Station (BAS) 2/6	Non-Medical Ground	**34**	Immediate	Blunt	Pt brought to Fox Co Firm Base after IED exploded 4-6 feet from him. Marines on scene state he was thrown 10-15 feet from force of blast. Pt was lethargic and …	Surgical Company Charlie	Urgent
1112	Forward Resuscitative Surgical System (FRSS)1/Shock Trauma Platoon (STP)2	CASEVAC	**57**	Immediate	Amputation	Probable RPG round through right upper quadrant (RUQ) abdomen, chest, and arm. Initially treated at BAS with valve dressing over chest wound and field dressing and tourniquet for abdomen and arm …	DOW	
1113	FRSS3/STP7	Medical Air	**13**	Immediate	Penetrating	20 y/o ad USMC, IED rear passenger HMWVV. C/O B/L leg pain, no LOC, …	Evac	Urgent
1114	BAS 2/6	CASEVAC	**9**	Immediate	Perforating	S. Pt. suffered GSW to L aspect of neck. Wound approx 5cm just inferior to mandibular angle, …	Evac	Priority
1115	Surgical Company Charlie	CASEVAC	**26**	Immediate	Hemothorax	18 y/o Marine - driver of High-backed Humvee struck by blast of IED. Occurred 30min prior to arrival. Per Corpsman …	DOW	
(Many more entries)								

Next we discuss our research approach to determine patient records in the CTR EMED that would apply to a statistical modeling effort. Appendix B contains an overview of our approach and shows our overall objective and how, for reasons to be presented, we divided the research effort into two major phases.

Algorithm Research for Life Threatening Records Identification

Because population of data fields in the CTR EMED was ongoing when this research began, we made use in an early development effort of an NHRC Excel file containing a small number of injuries that had resulted in killed in action (KIA) or DOW events (Galarneau, 2005). Each record in this file contained the necessary data points for our analysis.

The first algorithm developed (shown in the "early research" path in Appendix B) took data from the KIA/DOW file and augmented it with available CTR EMED records to capture the three PC codes that had descriptions that most closely matched the injury description in each record. Then, we augmented this data set with newly available CTR EMED records. The algorithm was then able to match only one of the file designations in 24% of the cases. A second algorithm version matched the PC code grouping in 82% of the cases, but it would subsequently require a means (automatic or manual) to pull out the correct PC. It was eventually decided to forego the effort to automatically select a PC code and investigate the possibility of identifying an LT risk category instead.

Approximately 2,500 CTR EMED records were available for this analysis. The large majority of the records corresponded to interventions at an FRSS or surgical company (SC). Records with missing timing data (for time of injury and MTF entry/exit) was supplied by SMEs with OIF medical deployment experience. This allowed us to maximize the number of "complete" records for the statistical analysis. Then we concentrated on identifying each injury as LT or not. Finally, we estimated the degree of mortality risk [(H), (M), or (L)]. Based on several passes, and results of technical interchange meetings (TIMs), the LT assignment was improved to be within 78% agreement with SME close analysis judgments. The CTR EMED Data Analysis (DA) Tool aided the SMEs in quickly sifting through records that had been estimated as LT by our algorithm. Eventually we found that 81 (H) risk casualties at an FRSS or SC had sufficient timing data for an initial survival analysis. Further improvements to the algorithm could have boosted the agreement rate, but we decided to switch away from the algorithm development effort and concentrate on a newly available CTR EMED file that contained a large number of ISS-scored records, a good measure of mortality (NHRC/Teledyne Brown Engineering [TBE] TIM, April 24, 2007). (See Appendix C for a more detailed chronology.)

Using Injury Severity Score as a Metric for Mortality Risk

Let us refer to the first two boxes in the "later research" path of Appendix B. In using the ISS values we assumed that values between 9 and 14 corresponded to a low risk of mortality, values between 15 and 24 corresponded to a medium risk, and lastly, values of 25 and higher corresponded to a high risk. The augmented data file with ISS values of 9 or higher was slightly over 1,000 records. Of these, approximately 80% corresponded to casualty records at Level IIa MTF (surgical) facilities as follows: Ar Ramadi Surge (6.7%), FRSS1/STP2 (36.9%), FRSS2/STP4 (9.7%), FRSS3/STP7 (3.0%), SC Alpha (14.3%), SC Bravo (11.8%), and SC Charlie (17.6%). The remaining records (approximately 20%) were from several MTFs (BAS, STP, and Landstuhl Regional Medical Center), none of which, taken separately, represented a sample large enough for a statistical analysis.

Too few Level IIa MTF arrivals had complete information on time of wounding or any record of prior treatment to permit a reasonable statistical analysis of mortality effects due to specific paths or delays in arrival (one of our original goals). However, the Level IIa MTF casualties with complete timing information did allow an investigation of the random variable for time to death in a Level IIa MTF facility. The effects of delays in arriving at an MTF and the subsequent treatment timing there are both important in modeling casualty mortality. Additional records may allow a more definitive empirical study of the effects of treatment delays in the future. An effort here to model the random variable for time to death within a Level IIa MTF is certainly worthwhile in the interim.

Some 160 records with ISS values ≥25 (for the subset of (H) risk casualties with sufficient timing information on arrivals and departures) were available across all Level IIa MTF facilities. (Although not attempted here, the single FRSS1/STP2 Level IIa MTF facility, a fairly large, homogenous group, would be an interesting analysis.) Of the 160 group, 17.1% were labeled DOWs at the Level IIa MTF, and the remainder was labeled as evacuations to the next higher MTF. Some 38% of the records from this group of 160 required significant SME input to provide casualty timing information for either arrival or departure. Corresponding groups of medium (M) and low (L) risk casualties contained <40 cases and were judged too small for consideration.

Statistical analyses of these ISS data are presented in the next section using nonparametric and parametric methods from the biomedical sciences and applied life data analysis literature.

Modeling Results for Injury Severity Score-Derived Level IIa MTF High-Risk Records

This section presents a statistical analysis of the ISS data results (Mitchell, 2006; NHRC/TBE TIMs, 24 April 2007, 13 March 2008). Figure 3 is a plot showing the times, from Level IIa MTF entry, to DOW or evacuation for the 160 high (H) risk cases. Each bar represents an individual patient's disposition. The evacuations represent only partial information on survival since these casualties were removed "alive" and no further information was available regarding their subsequent disposition. In classical biomedical science studies of this nature, these data are said to be censored or incomplete.

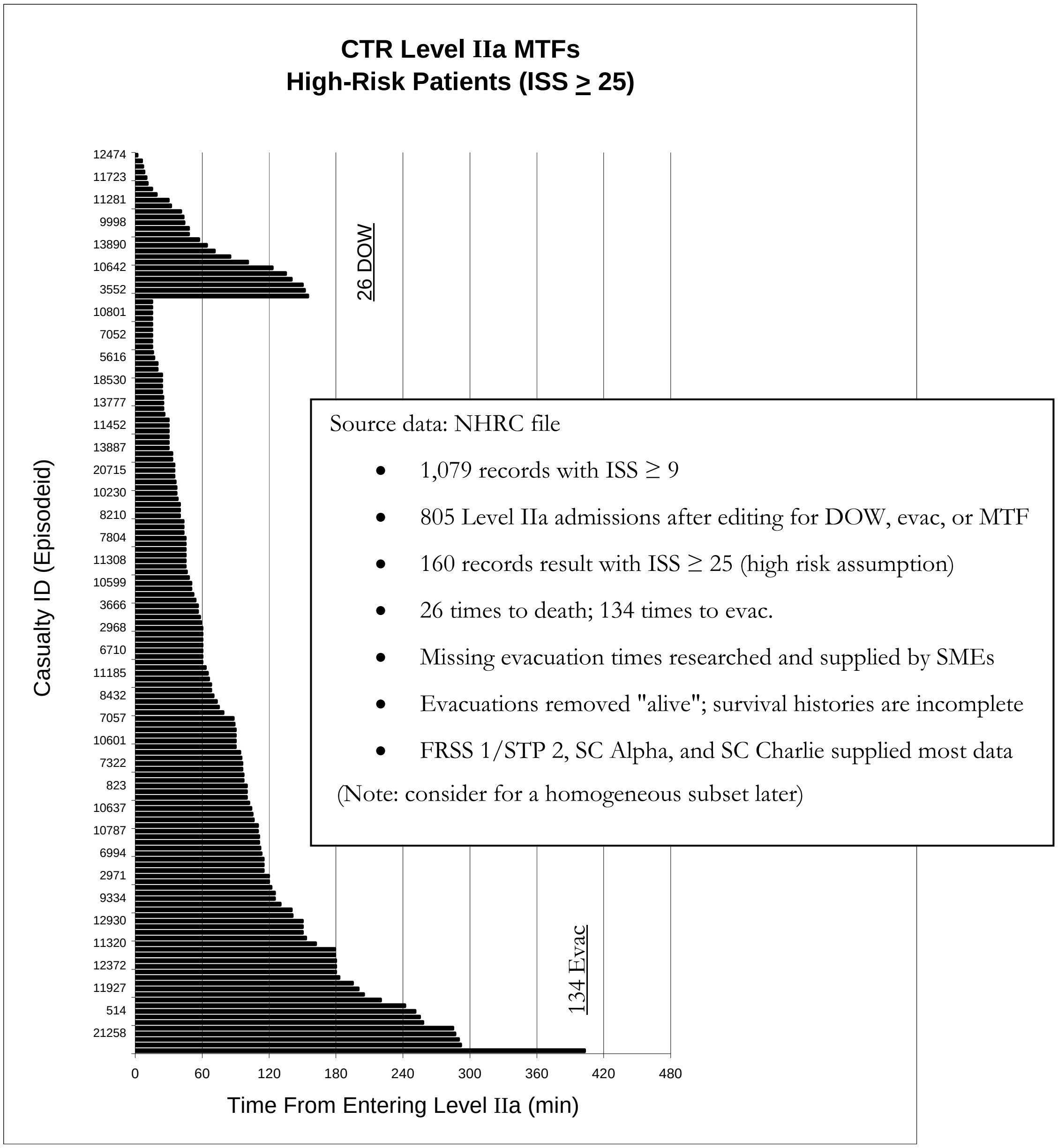
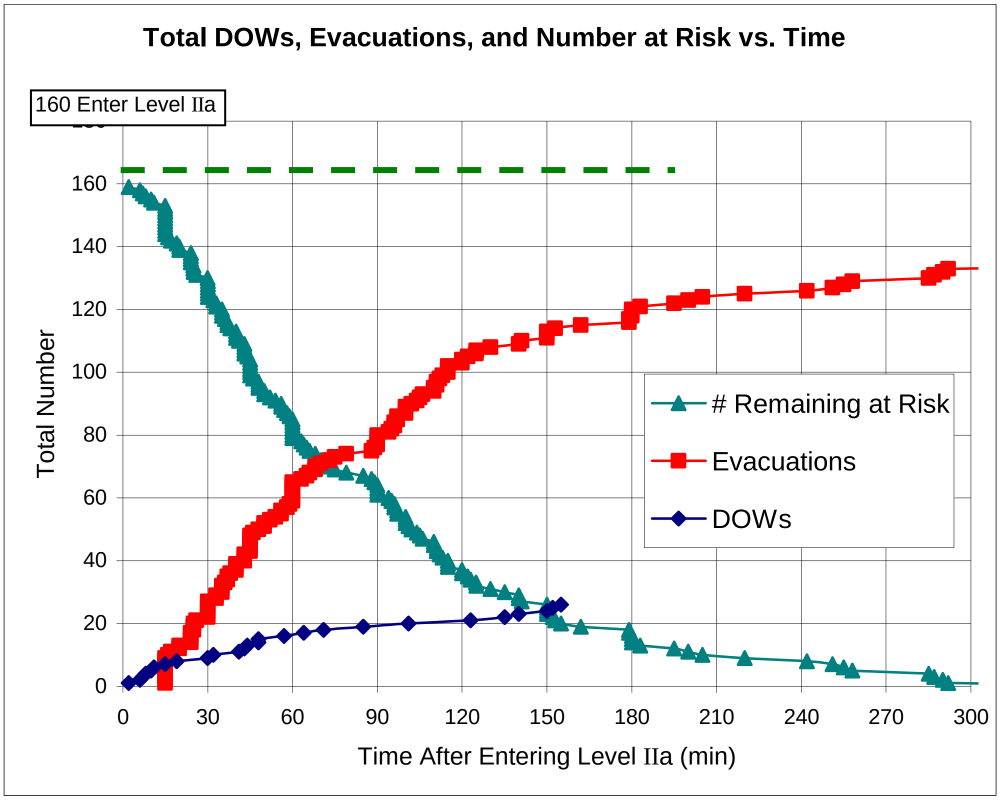


Figure 3. Duration in Level IIa MTF for deaths and evacuations of (H) risk casualties.

A rearrangement of the data used in Figure 3, to show a time-ordered plot of the total DOWs and evacuations, is given in Figure 4. It also shows the number of Level IIa MTF casualties remaining at risk versus time. This is computed by subtracting the total number of DOWs and evacuations from the entry number (160). These are called progressively censored samples in the literature of life data analysis. This practice differentiates casualties who enter and leave at different times, with deaths being intermixed with transfers (casualty is alive) (Gross & Clark, 1975; Nelson, 1982). Throughout, the random life-time T is assumed to have a PDF of f(t) and a probability of surviving past time t given by the survival function S(t) = Pr[T ≥ t]. Appendix D shows these functions and several other functions of interest for the Weibull distribution (Nelson, 1982). We concentrated initially on this statistical model in order to parallel the MD SME panel findings.



1) ISS ≥ 25 is an estimated high risk of mortality

2) Overall, 26 of 160 patients (16.2%) died in the various Level IIa MTFs.

3) Due to missing times in CTR EMED files, 38% of the evacuation times estimated by SME

4) There appear to be clumps of evacuations at 15, 30, ..., 180 min. (perhaps these were scheduled events)

5) Various methods from the biomedical sciences are available to estimate the empirical hazard and survival functions for data of this nature

Figure 4. Time histories for total DOWs, evacuations, and casualties at risk (n = 160, Level IIa MTF case).

We assumed that the casualties in the CTR EMED file entering the Level IIa MTFs are a random sample from the OIF severely injured population, and that the injuries incurred are a random sample of the battlefield wounds likely to be received during combat in that theater of operations. It is further assumed that the majority of casualties arrived at Level IIa MTFs immediately after first responder treatment with only a nominal delay. Our basic approach became characterizing the PDF of T by exploiting various features of the Weibull distribution. Both distribution and distribution-free techniques are used.

In the next section, survival function S(t), and hazard function h(t), each giving a different view or interpretation of the mortality process, are estimated from data plots and analytical means. The lengths of life times, via S(t) and the rate at which deaths occur (given by the hazard function), are different views of the same process.

Weibull Parameters via Graphical and Maximum Likelihood Methods

Before beginning a discussion of Weibull parameters derived from graphical and maximum likelihood methods of estimation, note that the data presented here are said to be "right censored" in that the dependent variable (time of death) for evacuated casualties is greater than the time of evacuation, and the actual value is unknown. These are often called incomplete samples (Elandt-Johnson & Johnson, 1980). We assume that if these casualties had remained in the Level IIa MTF, the subsequent time of death probability would approximate the distribution observed earlier for the DOWs. That is, we assume the censoring process is noninformative of future survivability. This is a common assumption in the literature for graphical and analytical techniques applied to life data (Nelson, 1982).

Probability theory provides essential tools for survival analysis of this type of clinical data. Possibilities for departure from theoretical models are many, and considerable detail is typically required for fitting distributions to observations. Graphical methods are usually emphasized to complement analytical methods such as maximum likelihood. Each method provides additional information (Nelson, 1982). We include both analytical and graphical methods in our approach to examining the several data files from the CTR EMED for OIF injuries.

Various nonparametric techniques are available to estimate the empirical S(t). They range from a so-called actuarial method dating back 200 years (and used to construct human life tables for insurance company risk analysis), to the modern use of individual ordered observations in continuous time. Each technique is used to estimate the probability of survival past some point in time. Estimates of the hazard function and the standard error of S(t) are usually provided by each technique. Example results of these methods were provided in various NHRC/TBE TIMs. The Kaplan-Meier method in continuous time is used here for a nonparametric estimate of S(t) and related functions (Elandt-Johnson & Johnson, 1980; Gross & Clark, 1975; Nelson, 1982). We will also use the analytical method of maximum likelihood assuming a Weibull distribution of life times to estimate the stochastic mortality process within the Level IIa MTFs.

The basic idea presented by Kaplan-Meier is to order the intermixed death and evacuation times and compute S(t) after each death using the recursive formula (8) (Appendix D). If there are assumed to be n event times in the sample, they are ordered from smallest to largest and numbered backward with reverse ranks. Table 2 illustrates the basic approach by showing the DOW/evacuation events for the initial 20 minutes within Level IIa MTFs where evacuation times are labeled "0" indicating censoring or incomplete data and "1s" indicate death events where the timing is known. Estimates of S(t) are computed using Excel and shown in the table. Table 3 shows the results for the entire data set using SYSTAT software (SYSTAT, 2004). Again, we point out that estimates of survivability are only possible at the time of death, so no results are shown past 155 minutes. Table 3 also shows various mortality metrics indexed by the death times. The cumulative hazard function and Weibull entries are discussed next.

Table 2. Example Data: Kaplan-Meier Estimate of S(t) for First 20 Min

Event Time t, min	Event	0-censor 1-complete	Nbr Remain at t, r_i	Kaplan-Meier Est of S(t)
2	DOW	1	160	0.994
6	DOW	1	159	0.988
7	DOW	1	158	0.981
8	DOW	1	157	0.975
10	DOW	1	156	0.969
11	DOW	1	155	0.963
15	DOW	1	154	0.956
15	Evac	0	153	0.956
15	Evac	0	152	0.956
15	Evac	0	151	0.956
15	Evac	0	150	0.956
15	Evac	0	149	0.956
15	Evac	0	148	0.956
15	Evac	0	147	0.956
15	Evac	0	146	0.956
15	Evac	0	145	0.956
16	Evac	0	144	0.956
17	Evac	0	143	0.956
19	DOW	1	142	0.950
20	Evac	0	141	0.950
...				

The equation for the cumulative hazard function of the Weibull distribution in equation (5) (Appendix D) can be transformed into equation (6) to give a linear relationship. This can be exploited to estimate the parameters {a, b} if the Weibull distribution applies to these data. So if the log10 of time t, versus the log10 of the cumulative H(t) plots a straight line, then a simple graphic results to visually gauge the viability of the Weibull distribution. Linear regression can be used to estimate the slope and intercept coefficients in equation (6) by the technique of least-squares analysis to give quantitative estimates of the straight line coefficients (Draper & Smith, 1966). Figure 5 shows the ordered pairs [log (H(t), log (t)] computed from Table 3 entries and the fitted linear regression model. In Figure 5, H(t) is estimated by summing the nonparametric estimates of the failure rate, h(t) = 1/k, where k is the number remaining at risk at time t.

Table 3. SYSTAT Used for Kaplan-Meier Estimate of S(t) and Other Measures

Death Time (t, min)	Number at Risk	Number Dying	K-M Est of S(t)	Std Error of S(t)	Hazard Rate, h(t)	Cumulative Hazard, H(t)	Weibull Model Survival Prob (via MLE)	Model Hazard Rate
2	160	1	0.994	0.006	0.006	0.006	0.995	0.002
6	159	1	0.987	0.009	0.006	0.013	0.986	0.002
7	158	1	0.981	0.011	0.006	0.019	0.983	0.002
8	157	1	0.975	0.012	0.006	0.025	0.981	0.002
10	156	1	0.969	0.014	0.006	0.032	0.977	0.002
11	155	1	0.962	0.015	0.006	0.038	0.975	0.002
15	154	1	0.956	0.016	0.006	0.045	0.967	0.002
19	142	1	0.95	0.017	0.007	0.052	0.959	0.002
30	131	1	0.942	0.019	0.008	0.059	0.938	0.002
32	124	1	0.935	0.02	0.008	0.067	0.935	0.002
41	111	1	0.926	0.022	0.009	0.076	0.919	0.002
43	110	1	0.918	0.023	0.009	0.085	0.915	0.002
44	106	1	0.909	0.024	0.009	0.095	0.914	0.002
48	98	2	0.891	0.027	0.010	0.105	0.907	0.002
57	89	1	0.881	0.029	0.011	0.116	0.892	0.002
64	78	1	0.869	0.03	0.013	0.129	0.88	0.002
71	72	1	0.857	0.032	0.014	0.143	0.869	0.002
85	68	1	0.845	0.034	0.015	0.158	0.848	0.002
101	52	1	0.828	0.037	0.019	0.177	0.824	0.002
123	35	1	0.805	0.043	0.029	0.206	0.793	0.002
135	31	1	0.779	0.049	0.032	0.238	0.777	0.002
140	30	1	0.753	0.054	0.033	0.271	0.77	0.002
150	27	1	0.725	0.058	0.037	0.308	0.757	0.002
152	23	1	0.693	0.064	0.043	0.352	0.754	0.002
155	21	1	0.66	0.069	0.048	0.399	0.751	0.002

Obviously the Weibull distribution describes these Level IIa MTF data well as indicated by the straight linear fit and near perfect correlation value shown in Figure 5. Least squares estimates can be transformed as indicated in equation (6) to give the Weibull parameter estimates.

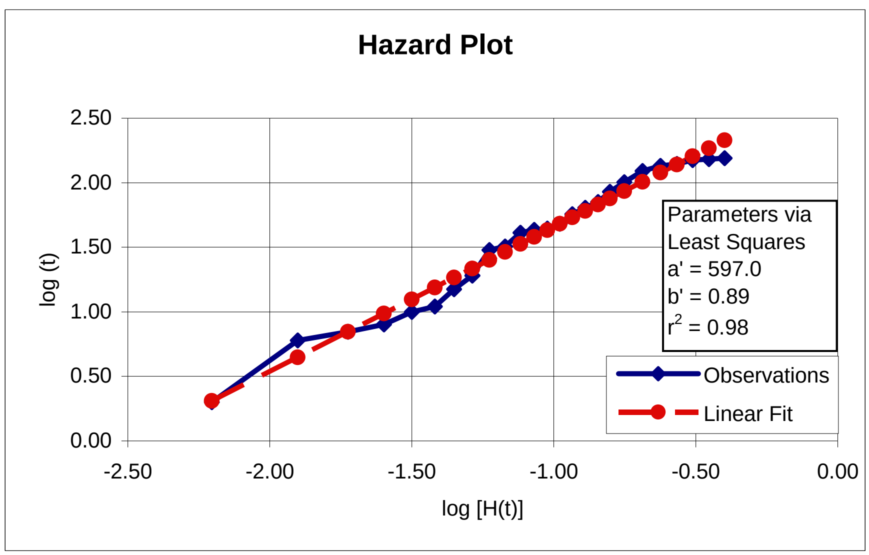


Figure 5. Log hazard plot and linear regression estimates of Weibull {a, b}.

Next we examined the maximum likelihood estimates of {a, b} and compared the parametric model based on the Weibull PDF with the nonparametric Kaplan-Meier estimates given in Table 3. Approximate confidence intervals are presented for parameters and S(t).

There are many excellent references available that give the details of parameter estimation by the method of maximum likelihood (i.e., locate the parameters needed to optimize the likelihood function of the data). The explanations given in Miller (1981) and Nelson (1982) are particularly instructive and are our preferred references. The SYSTAT software package (SYSTAT, 2004) is used to provide computational values (such as point estimates of the parameters) and their covariance matrix.

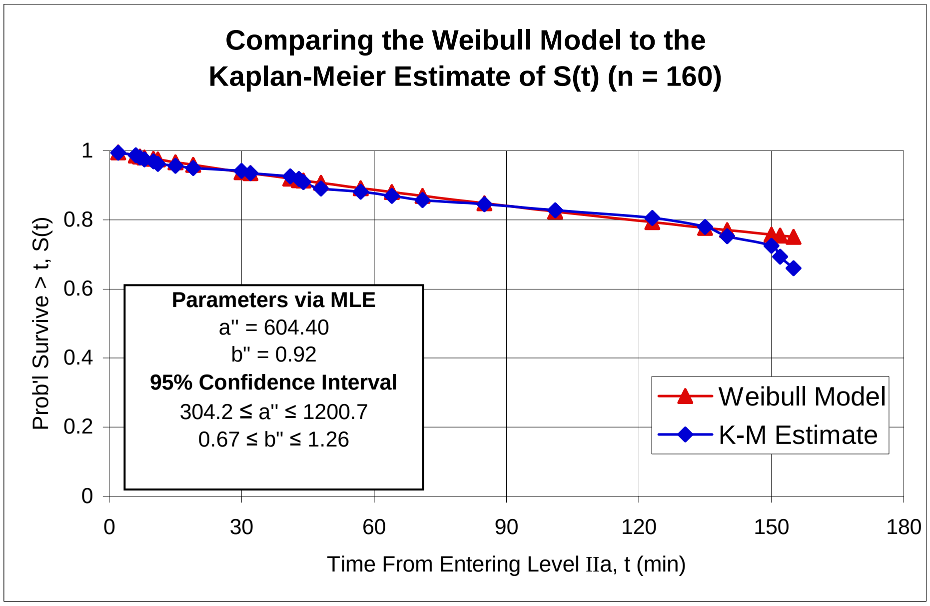


Figure 6. Nonparametric (Kaplan-Meier) and parametric (Weibull) estimates of S(t).

Figure 6 shows the Kaplan-Meier nonparametric estimate of S(t) and also the estimate based on the parametric Weibull model (both from Table 3). The closeness of these two estimated curves is encouraging because it confirms how applicable the Weibull PDF is to Level IIa MTF (H) risk casualty data.

The confidence intervals for the Weibull parameter estimates appear wide when taken independently, as shown in Figure 6. Because the sample correlation value from the SYSTAT run between the parameter estimates was substantially negative (-0.79) for the 160 observations, a joint confidence region for the two parameters would be a better way to depict variations about the point estimates (Draper, 1966; Gross & Clark, 1975). Work continues to plot the likelihood function for this sample data and compute the joint confidence region. This may lead to a better understanding of the characteristics of the sample mortality results.

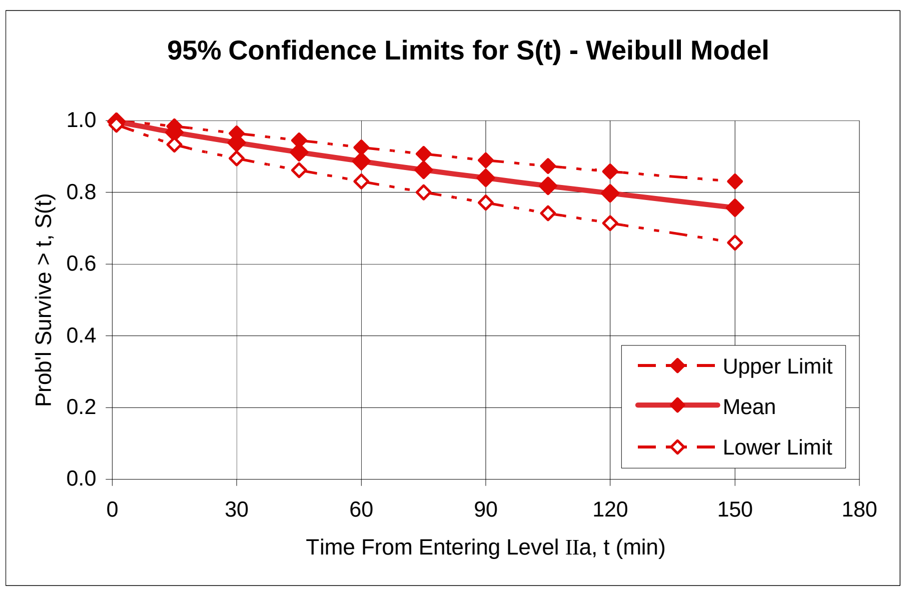


Figure 7. Approximate 95% confidence limits for Weibull survival function S(t).

Figure 7 shows the 95% confidence limits about the Weibull S(t) estimate. The probability of survival past 60 minutes is expected to lie between approximately 0.83 and 0.92 where the mean estimate is about 0.88. (We note that Nelson's [1982] methodology of computing approximate confidence intervals, by assuming a "ln (loge) normal" distribution for the sample parameter estimates, is used here. This is because SYSTAT does not provide special treatment for strictly positive parameters such as "a" and "b" in the Weibull distribution.)

Discussion

Comparing CTR EMED-Based Modeling Estimates with MD SME Results

The two plots labeled "MD SME…" in Figure 8 come from the original data collection effort (2003) for an (H) risk designated PC (such as PC 087), and the casualty flow case "first responder to STP/FRSS" (Mitchell, et al., 2004). These plots show how the panel expected delay times from POI to Level IIa MTF entry to influence mortality. (Note that the panel was conducted in a Delphi-like manner [Aylward, 2004].) From the CTR EMED file, times from POI to STP/FRSS entry averaged 49 minutes in 90 records that had complete POI to Level IIa MTF entry timing. We believe that the large majority of these records correspond with the MD SME cases with no intervening treatment after self/buddy aid (i.e., no routing through a BAS before Level IIa MTF entry). Presumably the upper MD SME curve for a 30-minute delay would shift down in some proportional manner for the empirical delay results observed here. No attempt is made to estimate that shift. The Weibull curves look very good in basic shape and location when compared with the MD SME results,

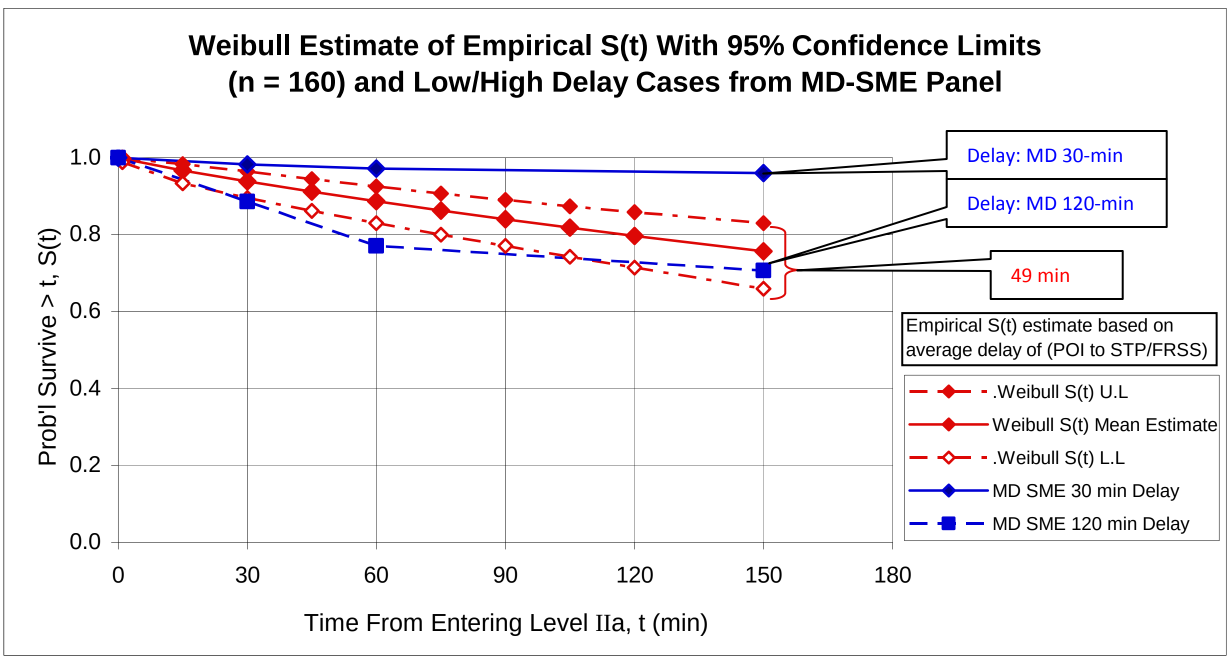


Figure 8. Weibull estimate of S(t) with confidence limits.

and we believe this justifies continued use of the Weibull model in TML+. The results are particularly impressive for times less than 60 minutes. (Note that these results are limited to (H) risk casualties where the CTR EMED ISS score is 25 or higher; PC 087 is an example.)

Some caveats are worth mentioning:

- While these analyses appear to support the Weibull model for survivability estimates (Figure 8), we did find a considerable amount of missing data points for exit/entry times. These required estimation by NHRC SMEs with OIF experience in the medical treatment and evacuation processes. But, as previously mentioned, data errors and omissions are fairly common in life data analysis (Elandt-Johnson & Johnson, 1980). These missing data points were supplied before any model fitting began.

- The MD SME results shown here is an average of the responses by nine MDs. The early results (out to 60 minutes) track very well. The variation in their individual responses does not appear here, but the standard deviation of responses at 60 minutes was 0.030 for the 30-minute delay curve, and 0.028 for the 120-minute delay curve. These are very small errors around the mean at this reference point, and this indicates a very consistent response by the panel. However, even a response of 2 standard deviations on the lower side of the upper 30 minute curve would approach the Weibull curves, and any adjustment for the 30- to 49- minute delay difference would further improve the comparison. Starting at 60 minutes, the Weibull curves tend toward the higher delay curve (120 minutes).

Summary and Future Activities

This project had two main objectives. First, we set out to describe a research effort to data mine the CTR EMED database and to identify individual PCs and the associated risks of mortality. The main problems with CTR EMED data mining and algorithm development had been inconsistencies in injury descriptions, and the labor required by SMEs to sift through the records for PC or LT classification. An algorithm was developed anyway, and together with the CTR EMED DA Tool, the algorithm facilitated SME navigation through the vast CTR EMED database. Then this effort was tabled when a large number of ISS became available in the CTR EMED file. The findings described here are encouraging. A detailed account of this research effort is included in Appendix B and C for future reference; there is more work that can be done.

Secondly we wanted to determine if a probability model could adequately describe the mortality events in a casualty's medical treatment flow. This effort was directly related to confirming the efficacy of the mortality modeling approach used in TML+. Considerable effort was spent completing records with missing timing data points for POI or entry/exit from MTFs. Of the records with an ISS of 25 or higher, 160 records had both entry and exit timing information at a Level IIa MTF and were judged adequate for a statistical analysis. Because data regarding the exact timing and location of injury were often missing, the effects of delay before reaching a Level IIa MTF facility was not addressed. Only the one path (first responder to Level IIa MTF) was deemed adequate to consider.

The results obtained were encouraging. The survival probability curve obtained [with approximate 95% confidence limits in Level IIa MTFs for (H) risk casualties] showed a strong graphical agreement with results estimated by the 2003 MD SME panel. Parameter estimates for the Weibull distribution were made using the method of maximum likelihood. A more detailed analysis of a joint confidence region for the parameters is now in progress because there was a fairly large negative correlation observed between the parameters in the sample data.

Our results (while applicable to just one of the MTFs in TML+) are very encouraging for two reasons: (1) they strongly reinforce the continued use of the Weibull PDF for use in TML+ in Level IIa MTFs, and (2) they confirm the effectiveness of NHRC's special panels to provide expert opinions when empirical results are unavailable.

Because the comparison of empirical and MD SME results is so noteworthy in the case presented here, it seems to confirm the legitimacy of continuing to use the MD SME inferred Weibull modeling results for the other MTFs and mortality risk categories used in TML+. When more data are available for MTFs closer to the POI and other risk categories, the Weibull model is a strong candidate for consideration in a future statistical analysis methodology. Looking back at Figure 1 again, we are intrigued to know what having early mortality data (before Level IIa MTFs) would reveal. These steeper curves are where casualty mortality is more dramatically influenced.

Other items that would be interesting to research include:

- The results that were analyzed combine all Level IIa MTFs. An interesting side study would be to examine the FRSS1/STP2 set that comprised about 37% of the data that we examined. Because this data set is presumably more homogenous, additional insight into the stochastic model for mortality may be possible. It might also be worthwhile to apply the data-mining LT algorithm to all of the CTR EMED records for this subset.

- Any delay effect (on the Weibull) from the POI results at the Level IIa MTFs was not examined since too few cases would have been available due to missing data points for the time of injury. It seems important to obtain more empirical data to look at this important concomitant variable for the Weibull model in Level IIa MTFs.

- These results were obtained from records of wounding from OIF in the CTR EMED. It would certainly be interesting to make a similar study using Operation Enduring Freedom injury data, if sufficient records are available.

These suggestions for future research would be logical next steps in developing a more in-depth understanding of battlefield mortality probability modeling for tactical medical logistics studies and analysis, and decision making.

References

Aylward, M. (2004, June 23). *Rapid Prototyping of Wireless Tools for Polling Groups, A Case Study*. Presentation at 72nd MORSS, Naval Post Graduate School, Monterey, CA.

Draper, N., and Smith, H. (1966). *Applied Regression Analysis*. Wiley.

Elandt-Johnson, R. and Johnson, N. (1980). *Survival Models and Data Analysis*. Wiley.

Galarneau, M. (2005, September). "DOW_KIA Mortality Curves," Excel file, Naval Health Research Center (NHRC), San Diego, CA.

Galarneau, M., and Mitchell, R. (2003, November 1). *Determining Subject Matter Expert (SME) Based Mortality Curves for Use in NHRC's Tactical Medical Logistics (TML) Planning Tool*. Presentation to MD SME Panel.

Galarneau, M., and Mitchell, R. (2004, June 23). *Modeling Dynamic Casualty Mortality Curves in the Tactical Medical Logistics (TML+) Planning Tool.* Presentation at 72nd MORSS, Naval Post Graduate School, Monterey, CA.

Galarneau, M., Hancock, W., Konoske, P., et al. (2004). U.S. Navy-Marine Corps Combat Trauma Registry Operation Iraqi Freedom-1 Preliminary Findings. Technical Report TR No. 04-20, Naval Health Research Center, San Diego, CA.

Gross, A., and Clark, V. (1975). *Survival Distributions: Reliability Applications in the Biomedical Sciences*. Wiley.

Konoske, P. (2008, August 6). *Medical Modeling and Simulation: From Data to Planning Tools*. Briefing prepared for Joint Health Planners Conference, Newport News, VA.

Lowe, D., Hill, M., and Galarneau, M. (2008, March). "Review of Patient Codes and Associated Mortality Classifications as Used in the TML+ Planning Tool." Naval Health Research Center/Teledyne Brown Engineering, Inc. (TBE) technical notes.

Miller, R. Jr. (1981). *Survival Analysis*. Wiley.

Mitchell, R. (2006, November 21). "Interim Comments on Mortality Study Objectives and CTR Data Analysis." Teledyne Brown Engineering, Inc. e-mail to Naval Health Research Center.

Mitchell, R. (2007, March 15). "Some Statistics on the CTR file." Teledyne Brown Engineering, Inc. e-mail to Naval Health Research Center.

Mitchell, R., Galarneau, M., Hancock, B., and Lowe, D. (2004). Modeling Dynamic Casualty Mortality Curves in the Tactical Medical Logistics (TML+) Planning Tool. Technical Report No. 04-31. Naval Health Research Center, San Diego, CA.

Mitchell, R., and Galarneau, M. (2006). *Simulating Dynamic Mortality within the Military Medical Chain of Evacuation and Treatment.* Society of Computer Simulation, Spring Conference, Huntsville, Al.

Nelson, W. (1982). *Applied Life Data Analysis*. Wiley.

NHRC/TBE. (2006, June 6). Kick-Off Meeting: "DOW/CTR Analysis." Naval Health Research Center, San Diego, CA.

NHRC/TBE. (2006, December 12; 2007, April 24; 2008, March 13; 2008, April 15; 2008, October 7). Technical Interchange Meetings:."Developing a Statistically Based Model for Battlefield Casualty Mortality by OIF Empirical Patient Encounter Records from the Navy-Marine Corps Combat Trauma Registry (NMC CTR).". Naval Health Research Center, San Diego, CA.

NHRC. (2005, March 31). Proposal to ONR: "Developing a Dynamic Died-of-Wounds Statistical Model for TML+ using Navy-Marine Corps Combat Trauma Registry Data."

Pang, G. (2008, March 20). Combat Trauma Registry Data File: "TBE_080320_f1." Naval Health Research Center, San Diego, CA.

Parker, J. (2007, April 11). DOW PC Estimator Algorithm Summary with CTR Data Life-Threatening Criteria. Teledyne Brown Engineering, Inc. report, Huntsville, AL.

Parker, J.. (2006, October 11). CTR Data DOW PC Estimator Algorithm. Teledyne Brown Engineering, Inc. report, Huntsville, AL.

SYSTAT, Statistics III, www.systat.com, SYSTAT Software, Inc., Richmond, CA 2004.

Teledyne Brown Engineering. (2008, August 14). "ONR Annual Report on DOW Curve." Teledyne Brown Engineering, Inc. document submitted to Naval Health Research Center, San Diego, CA.

Appendix A. CTR EMED Theater Medical Registry Form

Navy-Marine Corps CTR – Theater Medical Registry Form

Name (Last, First MI):	Patient I.D. / SSN: - -	Rank:	Unit:

Date of Birth: DDMMMYY	Gender: ☐ Male ☐ Female	Blood Type:	Allergies:

MTF Patient Evacuated From: (If casualty rec'd from point of injury, enter POI)	MTF Designation:	MTF Location:	Facility Type: ☐ Base-X ☐ GP ☐ CBPS ☐ Hard Bldg

Medical Visit: ☐ Battle Injury ☐ Disease ☐ Non-Battle Injury ☐ Dental (Routine) **Treatment:** ☐ Initial ☐ Follow-Up

Date/Time of Injury: DDMMMYY/TIME

Date/Time of Arrival: DDMMMYY/TIME

Transport Care To Facility: (From Point of Injury to 1° MTF)
☐ Casualty Evacuation (CasEvac)
☐ En Route Care (ERC)
☐ Non-Medical
Transit Duration Time: _____

Arrival Method:
☐ Walked ☐ Carried
☐ Med Evac Ground ☐ Non Med Evac Ground
☐ Med Evac Air ☐ Non Med Evac Air
☐ Train ☐ Water Boat
☐ Unknown ☐ Other: _____

Category:
☐ US Marine Corps ☐ SOF
☐ US Navy ☐ Civilian
☐ US Army ☐ Contractor
☐ US Air Force ☐ Combatant
☐ Host Nation Security ☐ NGO - _____
☐ TCN: _____ ☐ Other: _____
☐ Unknown ☐ None

Wounded By: ☐ Enemy ☐ Friendly ☐ Self Accident ☐ Civilian (Host Country) ☐ Self Non-Accident ☐ Sports/Recreation ☐ Training ☐ Unknown ☐ Other: _____ ☐ N/A

Mechanism of Injury:
☐ Aerial Bomb ☐ Aggravated R.O.M. ☐ Assault/Altercation ☐ Bite / Sting ☐ Blunt Trauma ☐ Building Collapse ☐ Burn ☐ Crush ☐ Drowning ☐ Electrical/Electrocution
☐ Fall ☐ Flying Debris ☐ Grenade ☐ GSW/Bullet ☐ Helicopter Crash ☐ Hot Object/Liquid ☐ IED ☐ VBIED (Vehicle Borne) ☐ Knife/Edge(Stab) ☐ Landmine
☐ Machinery/Equipment ☐ Mortar ☐ Motor Vehicle Accident ☐ Parachute Drop ☐ Pedestrian ☐ Rocket ☐ RPG ☐ Unexploded Ordnance ☐ Chemical ☐ Biological ☐ Radiation/Nuclear
☐ Environmental ☐ Other: _____ ☐ Unknown ☐ N/A

Triage Category: ☐ Immediate ☐ Delayed ☐ Minimal ☐ Expectant ☐ N/A

Glasgow Coma Scale (Circle each)

Eye Opening	Verbal Response	Motor Response
1-None	1-None	1-None
2-To pain	2-Incomp. sounds	2-Extend pain
3-To command	3-Inapprop. words	3-Flex to pain
4-Spontaneous	4-Confused	4-Withdraws
	5-Oriented	5-Localize pain
		6-Obeys

Glasgow Score _____ (Enter total number)

Personal Protective Equipment:	Worn	Not Worn	Struck	Penetrated
Helmet - Circle: USMC / ACH / AVN / CVC / MICH	☐	☐	☐	☐
Eyewear - Circle: Wiley-X / ESS Land / ESS NVG / SG-1 / SWDG / BLPS / UVEX XC / Other	☐	☐	☐	☐
Ear Protection - Circle: Combat Ear Plugs / Single Flange / Other Circle: XS / S / M / L / XL	L ☐ R ☐	L ☐ R ☐		
Neck Protector - Yoke and Throat	Y ☐ T ☐	Y ☐ T ☐	Y ☐ T ☐	Y ☐ T ☐
Flak Vest / IBA - Circle: XS / S / M / L / XL / XXL / XXXL / XXXXL	☐	☐	☐	☐
Ceramic Plates - Circle: XS / S / M / L / XL	F ☐ B ☐	F ☐ B ☐	F ☐ B ☐	F ☐ B ☐
Axillary / Deltoid / Upper Extremity	L ☐ R ☐	L ☐ R ☐	L ☐ R ☐	L ☐ R ☐
Groin Protector	☐	☐	☐	☐
Leg / Lower Extremity	L ☐ R ☐	L ☐ R ☐	L ☐ R ☐	L ☐ R ☐

Care Prior to Arrival:

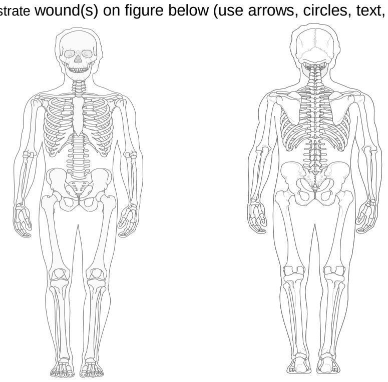
Tourniquet ☐ No ☐ Yes Type: _________ Time on: ______ Time off: ______
Airway ☐ No ☐ Yes Type: _______________
IV's ☐ No ☐ Yes Type: _____ Location: _____ Fluid: _________ Amount: ____ ml
C-Collar ☐ No ☐ Yes
Chest Tube ☐ No ☐ Yes ☐ L ☐ R ☐ Air ☐ Blood _____ ml
Needle Decompression ☐ No ☐ Yes ☐ L ☐ R ☐ Air ☐ Blood _____ ml
Temp Control Measures ☐ No ☐ Yes Type: ________________
Intraosseous Access ☐ No ☐ Yes Location: ________________

Vital Signs

Time	Temp	Pulse	Resp	B / P	SpO_2
				/	
				/	
				/	
				/	
				/	
				/	
				/	

AB	Abrasion
AMP	Amputation
AV	Avulsion
BI	Blast Injury
BL	Bleeding
Burn	Burn
C	Crepitus
Deform	Deformity
DG	Degloving
E	Ecchymosis
FB	Foreign Body
Frag	Fragment
Fx	Fracture
GSW	Gunshot Wound
H	Hematoma
Lac	Laceration
P	Pain
PW	Puncture Wound
SS	Seatbelt Sign
SW	Stab Wound

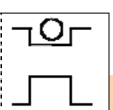
Illustrate wound(s) on figure below (use arrows, circles, text, etc.)



Current Treatment & Procedures

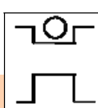
Oxygen	_____ L/min.
Fluid Administration: Crystalloid #1: _________ Crystalloid #2: _________ Colloid #1: _________ Colloid #2: _________ Other: _________	_____ ml _____ ml _____ ml _____ ml _____ ml / mg / gm
Intubated: In ____ Out ____	_____
CRIC	No / Yes
Sedated	_____
Chemically Paralyzed	_____
Needle Decompression ☐ Left: ☐ Blood ____ ml ☐ Right: ☐ Blood ____ ml	No / Yes
Chest Tube: ☐ Left: ☐ Blood ____ ml ☐ Right: ☐ Blood ____ ml	No / Yes
Intra-Osseous Access (Location)	__________
Foley Catheter	No / Yes
Collar / C-Spine	No / Yes
Tourniquet – – Time On: _____	Time Off: _____
Hemostatic Dressing (Type): _________	
Blood Products: _______	_____ Units/Pks
Auto Transfusion: No / Yes	_____ ml
Factor rFVIIa (NovoSeven)	_____ mg
Walking Blood Bank (FWB)	_____ Units
HBOC _______	_____ ml
Splints (Location)	_____

Pulses Present:
S=Strong
W=Weak
D=Doppler
A=Absent



Class of Hemorrhage: I ☐ II ☐ III ☐ IV ☐

*Highlight Burn Area – ☐ 1° ☐ 2° ☐ 3°
___ %Total Body Surface Area



DTR's:
0
0.5+ - 1+
2+ - 3+
4+ - 5+

Name (Last, First MI):

Prior Medical History:

☐ None ☐ Unknown
☐ Cardiac ☐ HTN
☐ Respiratory ☐ DM
☐ Seizure ☐ Ulcer (G.I.)
☐ Other: _____

Medications Administered:

Medication:	Dose #1 / Time	Dose #2 / Time	Dose #3 / Time
☐ Morphine	_____ / ____	_____ / ____	_____ / ____
☐ Versed	_____ / ____	_____ / ____	_____ / ____
☐ Fentanyl	_____ / ____	_____ / ____	_____ / ____
☐ Demerol	_____ / ____	_____ / ____	_____ / ____
☐ Ancef/Gent	_____ / ____	_____ / ____	_____ / ____

Subjective (Chief Complaint) / History:

Objective / Exam:

Laboratory Results:

X-Ray Results:

Assessment / Diagnosis: Primary -

Secondary -

Plan / Procedure(s) / Treatment / Operative Notes / Nursing Notes / etc:

DNBI Category: ☐ N/A
☐ Combat/Oper. Stress Rxn
☐ Dermatologic
☐ GI, Infectious
☐ Gynecologic
☐ Heat/Cold Injuries
☐ Injury, Recreational/Sports
☐ Injury, MVA
☐ Injury, Work/Training
☐ Injury, Other
☐ Ophthalmologic
☐ Psych., Mental Disorders

Disposition Date/Time:

DDMMMYY/TIME

☐ RTD
☐ Light Duty x _____ day(s)
☐ SIQ x _____ day(s)
☐ **Deceased - -** ☐ **DOW** ☐ **KIA (see below)**
☐ Up Chit (Aviation)
☐ Down Chit (Aviation)
☐ Evac to _____
☐ Admit to ____ x ____ day(s)
☐ Other: _____

Date/Time of Death: _____________ ☐ N/A

ANATOMIC:
☐ Airway ☐ Head ☐ Neck ☐ Chest ☐ Abdomen ☐ Pelvis ☐ Extremity (Upper/Lower) ☐ Other, specify: (_____________)

PHYSIOLOGIC:
☐ Breathing ☐ CNS ☐ Hemorrhage ☐ Total Body Disruption ☐ Sepsis ☐ Multi-organ Failure ☐ Other, specify: (_____________)

Provider Signature:

Provider Name (Printed or Typed):

Figure 9. Theater Medical Registry Form.

Appendix B. Overview of Data Collection and Analysis

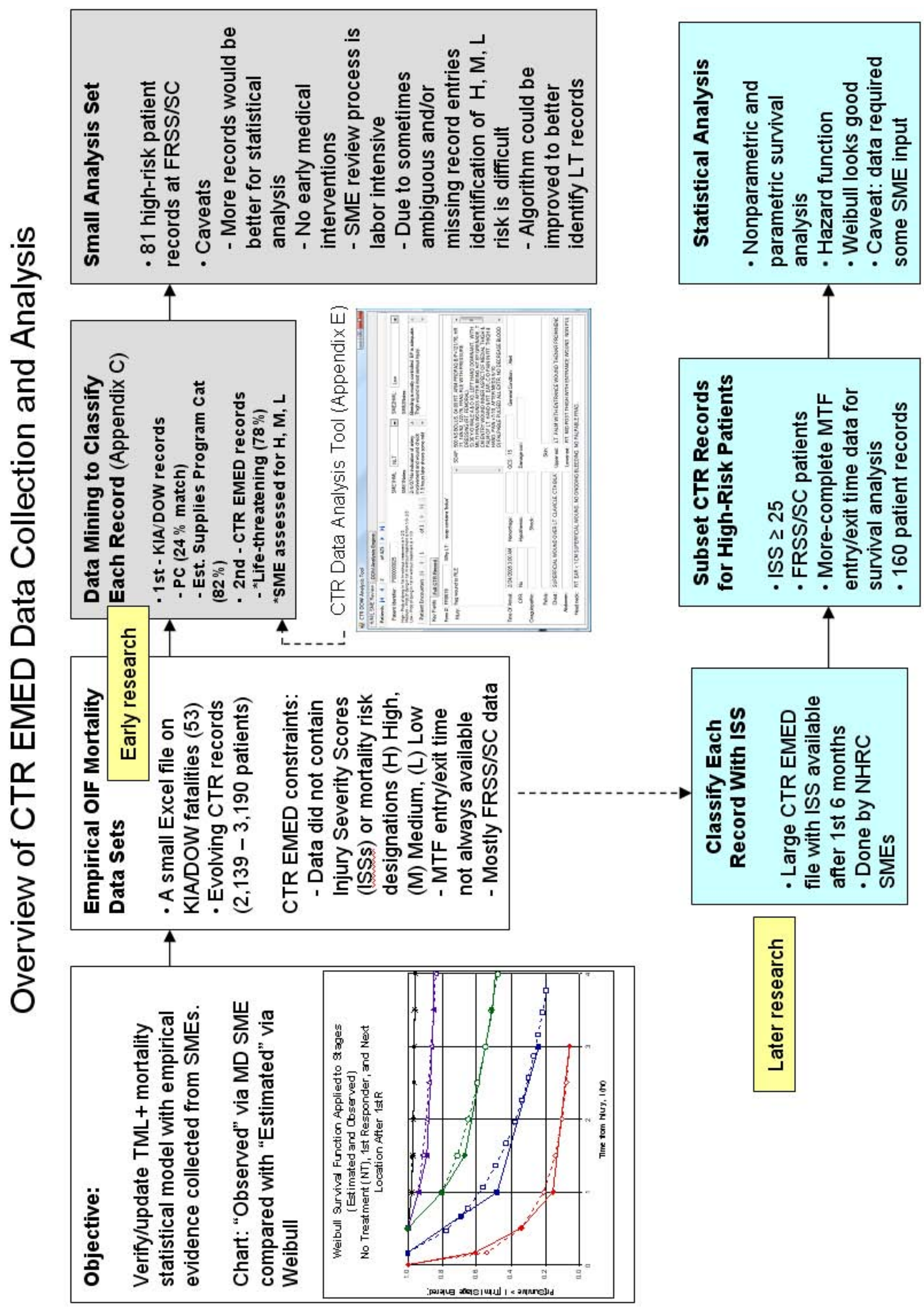


Figure 10. Overview of data collection and analysis.

Appendix C. Chronology of Algorithm Evolution

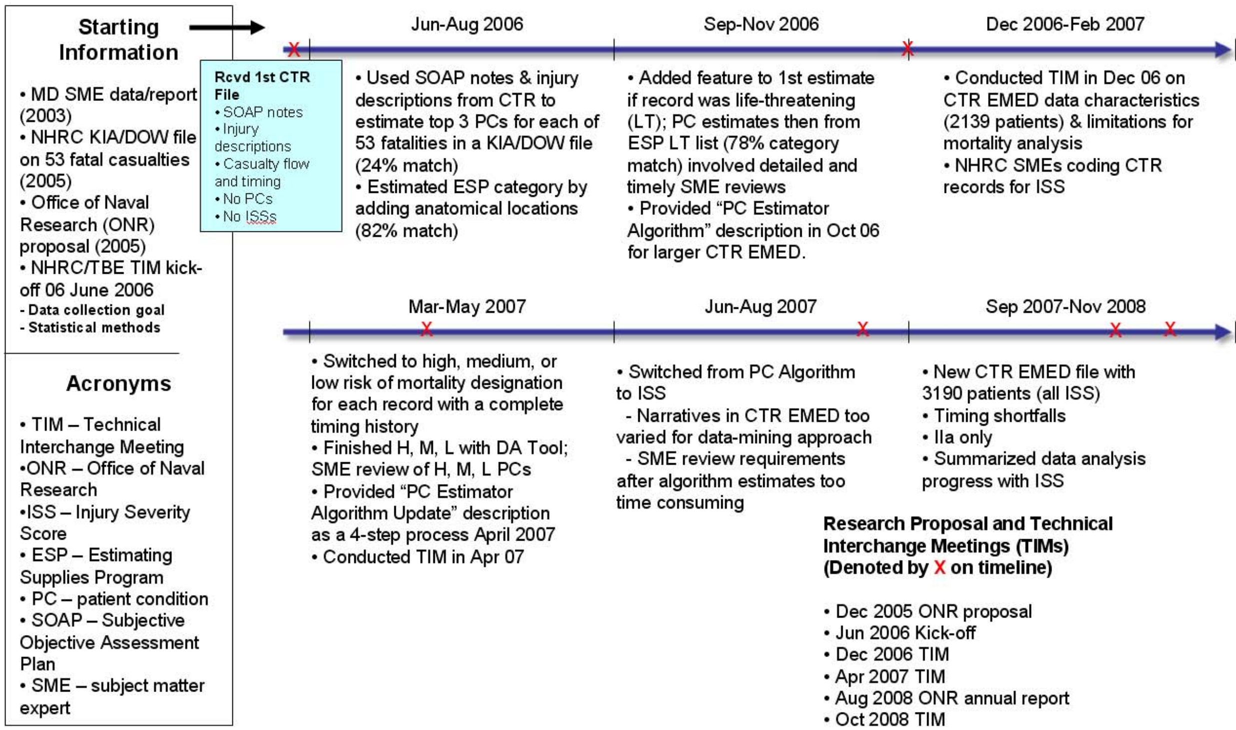


Figure 11. Algorithm timeline.

Background Notes

Our objective was to mine CTR EMED data (like MD SME PCs) for statistical analysis of battlefield casualties to determine mortality probabilities. The problem we found is that records are not coded by PC. The preferred approach would be to run a data-mining algorithm on CTR EMED files and pull out PCs that correspond to those used in the MD SME panel (e.g., PC 087).

A secondary approach, if a computer algorithm cannot be developed that can produce results with a high confidence, could be to manually (with SME assistance) determine if a record is LT and that it has an H, M, and L risk of mortality. Perform statistical analysis by using records in various risk categories. (This is the end product desired for the current TML+ application. Mortality is modeled at the risk category level.)

If this option is not workable, then another approach could be to determine if the records can be mined by first identifying the record's ESP category, and then (again, with SME manual assistance) find the PCs associated with that record. If the record indicates a casualty has LT injuries, the record can be flagged and set aside. After combing records in this way, the researcher can select the top 3 PCs from the approved LT list in the ESP category.

A tool available to the SMEs is the DA Tool. The SME can focus on the ISS to identify LT indicators. SMEs at NHRC add the ISS score to each CTR EMED record.

Information available to researchers as of July 2006:

1. MD SME data and report (Mitchell, Galarneau, Hancock, & Lowe, 2004; Mitchell & Galarneau, 2006)
2. NHRC KIA/DOW files containing 53 fatal casualties with PCs and ESP categories (Galarneau, 2005) (all confirmed as definitely LT)
3. ONR proposal of Dec 2005 (NHRC, 2005)
4. NHRC/TBE Kick-Off Meeting, TIM 6 Jun 06 (data collection goals and statistical methods established)

This algorithm research was a repetitive process of algorithm design, testing, and algorithm redesign. The final algorithm developed is described in Appendix E.

The NHRC/TBE team built several CTR EMED files containing SOAP notes, injury descriptions, casualty flow, timing information, and many other medical conditions for OIF patient records. ESP PCs and severity codes were not available at the start. ISS codes were added via manual effort, which began near the end of this period.

Figure 11 shows an annotated timeline for work on the algorithm and a list of various progress reports.

July–December 2006 Notes

This research kicked off in July, 2006. Initial algorithm development and testing began on an existing KIA/DOW Excel file. As more CRT files with a larger number of records became available, they were used.

The initial KIA/DOW file contained 53 records. Initially SOAP notes and various injury descriptions from these small CTR EMED records were used to identify the top three PCs for each fatality. Employing the algorithm, we obtained a 24% records match between one of the top three PCs and one of the PCs provided in the fatality file. Manually combing this file in the same way produced a 59% match. We were able to improve estimates of ESP categories by adding anatomical location entries to obtain an 82% match.

We eventually transitioned from the small KIA/DOW file to a large CTR EMED file of 2,645 records. On October 11, 2006, we had developed an initial PC Estimator Algorithm (Parker, 2006). On December 12, 2006, we conducted an NHRC/TBE TIM on CTR EMED data characteristics, algorithm results, and areas for improvement regarding mortality analysis applicable to TML+. Our data summary showed:

- missing timing data in the files (time of injury and/or MTF arrival/departure times)
- 2,139 records with timing information for analysis (both entry and exit from an MTF)
- few 1stR/BAS encounters were reported; most were FRSS or SC
- no ISS results (SME reviews and data coding were in progress)

Our algorithm summary showed:

- The algorithm estimated records to be “definitely LT,” then estimated ESP category, and finally selected top three LT PCs from chosen ESP category.

- Of 616 records estimated by the algorithm to be “definitely LT,” two ESP categories were used for algorithm evaluations.
- There were 37 abdomen and pelvis wound records and 141 multiple injury wounding (MIW) records.
- the outcome of labor-intensive SME manual review (as a confirmation of the algorithm’s ability to identify LT records) revealed that:
  - SMEs agreed with 58% of the algorithm’s designations (i.e., 103 of 178 records).
  - SME judgments agreed with the ESP category and one of the top three PCs chosen (29% of combined 178 test records).
  - SMEs agreed with the ESP category but none of PC choices 11% of the time (note that SMEs agreed with a combined 40% of the records on the ESP category choice by the algorithm).
  - SMEs disagreed with ESP category choice in 19% of the records.
  - SMEs thought records were either not LT or that the data were not adequate to determine a PC in 42% of the records.

We suggested areas where performance could be improved:

- Anatomical location data were often ambiguous and could be identified as superficial injuries (if this occurs, this will make the estimator appear to be in the wrong PC category).
- The estimator does not detect the difference between superficial and severe injuries.
- The estimator does not detect that treatments listed in SOAP notes indicate certain injuries/severities.
- The estimator does not detect medical visualization (for instance, gunshot wound to abdomen would have organ damage).
- LT is often erroneously assigned (mostly because of “immediate triage” callout in record).
- It was difficult for SMEs to extract the data they needed in the very large Excel file.

Considerations for further improving the algorithm included:

- Use ISS/Abbreviated Injury Scale scores to assist in determining severity
- Focus more on SOAP notes and injury descriptions; even though these are more difficult to interpret programmatically, they are what the SMEs used most in making their determinations
- Add improved system to select PC category based on injury description and SOAP notes
- Add more abbreviation synonyms
- Add some medical knowledge, associate treatments with injuries, and recognize potential injuries related to the primary injury
- Add adjective/noun pairing to associate words like “minor wound”
- Use severity and anatomical location to determine LT/non-LT (in addition to existing criteria, except for triage “immediate” category, which gave many false positives)

Considerations for improving the SME review process include:

- Create data entry application that shows Injury description, SOAP notes, and other relevant information in a cleaner, more organized way (to enable easier browsing and insertion of comments)

Recommendations for TIMs include:

- Address LT algorithm improvement above
- Improve the CTR EMED DA Tool for SME analysis of the abundant information in CTR EMED records
- Switch to an H, M, L risk of mortality designation for each record with complete timing information and improved LT "words"
- Begin effort to implement ISS as a surrogate for mortality (target expected 1,200 records to be completed by February 2007)

January–June 2007 notes

In this time period emphasis was placed on improving the algorithm. Based on the December 2006 TIM recommendations we tried to better identify LT records. The effort eventually transitioned to the use of a large number of ISS scores as a measure of mortality. This eliminated the need for a data-mining algorithm.

Algorithm improvements resulted in the improved identification of LT records (the set of 178 SME combed records stood as a control group; SME evaluation was considered the truth state). The outcome summary from the DA Tool helped immensely:

- The new algorithm yielded 39% fewer records designated as LT (108 abdomen and pelvis and MIW records vs. 178 before). This was shown to be an important improvement because the final evaluation to determine the validity of algorithm outputs would be the manual confirmation by SMEs. Here 78% of these were judged valid eliminations and 22% were judged to be incorrectly eliminated as non-LT records.
- The algorithm correctly identified 87 of the SME LT set as LT and correctly identified 55 of their non-LT/not enough data set for an overall statistic of 80% correctness. This was judged by us to be a good improvement at the LT record level.
- A more detailed analysis of SME involvements:
    - SMEs agreed with the ESP category and one of the top three PCs chosen (41% of the 178 test records)
    - SMEs agreed with the ESP category but none of PC choices (15% of the test records. Note that SMEs had by now agreed with a combined 56% of the records on the ESP category choice by the algorithm).
    - SMEs disagreed with ESP category choice (28% of the test records).
    - SMEs thought the record was either not LT or that the data were not adequate to determine a PC (17% of the test records).
- Provided "PC Estimator Algorithm Update" description as a 4-step process in April 2007 (Parker, 2007) (See Appendix E for a description of the DA Tool and LT criteria).
    - The above results were deemed unsatisfactory for automatically obtaining a large number of records for a statistical analysis.
- An attempt was made to use the LT records and an SME effort was conducted to identify H/M/L records from which a suitable number might be found to permit a statistical analysis to verify the mortality result in TML+.

- From a large OIF file (on 3,190 casualties) there were 436 casualties with automated LT indicators. (See TBE e-mail to NHRC on March 15, 2007 "Some Stats on the CTR file" [Mitchell]).
  - SME review and assessment agreed with 56% of these records being of an LT character (of these, the SMEs provided a breakout to H/M/L), but only 81 of these proved sufficient for an (H) risk comparison with TML.

- Conducted NHRC/TBE TIM April 2007 (NHRC/TBE, April 24, 2007)
  - Demonstrated new algorithm and DA Tool
  - Presented results from automated LT Algorithm and SME assessment
  - Discussed sample data on 81 (H) risk casualties showing DOWs and evacuations
  - Demonstrated Life Table format for analysis and preliminary analysis
  - Observations: no early data from POI in casualty stream; data are incomplete for the purpose of developing an algorithm to search for LT without SME support; decided to switch to ISS mortality values when a large file became available
    - Narratives in CTR EMED too varied for conclusive data-mining approach
    - SME review requirements to identify H/M/L records after algorithm estimates very labor intensive and time consuming
- Switched from PC Algorithm to use of ISS field now populated in CTR EMED:
  - ISS $\geq$9 assumed to be LT casualties
  - ISS $\geq$25 assumed to be (H) risk LT casualties

Appendix D. Features of Random Variable T Given by Weibull Distribution

(1) Probability Density Function, f(t), with parameters {a, b} f(t; a, b) = $\frac{b}{a}\left(\frac{t}{a}\right)^{b-1} e^{-\left(\frac{t}{a}\right)^{b}}$, a>0, b>0, t>0
(2) Cumulative Distribution Function, F(t) = Pr[T ≤ t] F(t) = 1- $e^{-\left(\frac{t}{a}\right)^{b}}$
(3) Survival Function, S(t) = Pr[T > t] = 1-F(t) S(t) = $e^{-\left(\frac{t}{a}\right)^{b}}$
(4) Hazard Function, h(t) = $\frac{f(t)}{S(t)}$, instantaneous failure rate at t h(t) = $\frac{b}{a}\left(\frac{t}{a}\right)^{b-1}$
(5) Cumulative Hazard Function H(t) = $\int h(s)ds$ from s = 0 to t H(t) = $\left(\frac{t}{a}\right)^{b}$
(6) Log-Log equation to estimate {a, b} from hazard plot log (t) = $\frac{1}{b}$ log H(t) + log (a)
(7) Basic relation between S(t) and H(t) S(t) = $e^{-H(t)}$
(8) Kaplan-Meier Estimate of S_i after each death event *i* at time t_i $S_i = \frac{r_i - 1}{r_i} S_{i-1}$, S_0 = 1, r_i is reverse rank of death *i*

Appendix E. Algorithm Description

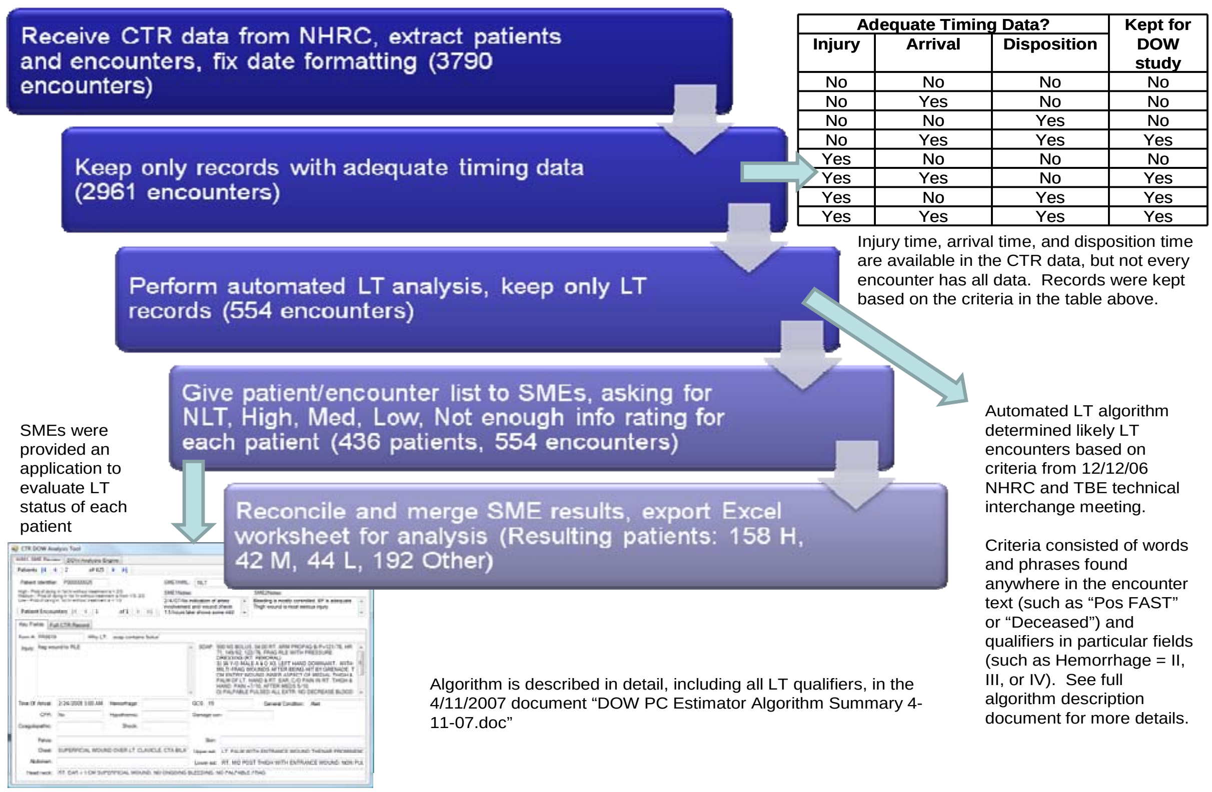


Figure 12. Extracting mortality related data from the CRT file.

CTR EMED Data LT Identification Criteria

The current method used in this study to determine the life-threatening (LT) casualties is outlined below.

1. The list of encounter records received from NHRC was split into two lists: casualties and encounters. Each casualty may have multiple encounters.
2. All dates and times were validated and combined into a single date/time field for easier computer analysis.
3. There are three date/time fields available (time of injury, time of arrival, time of disposition). Encounters that did not have adequate timing information for the DOW analysis were removed according to the table below.

Table 4. Decision Table for DOW Analysis

Adequate Timing Data?			Kept for DOW study
Injury	Arrival	Disposition	
No	Any	Any	No
Yes	No	No	No
Yes	Yes	No	Yes
Yes	No	Yes	Yes
Yes	Yes	Yes	Yes

4. All encounters were then subjected to an automated test for likely LT status, primarily based on key words and phrases established at the NHRC technical interchange meeting on 12 December 2006. Every encounter that was marked LT by the estimator was also marked with an indication of what caused the LT status. See Appendix F for a complete list of LT indicators used by the automated LT estimator.

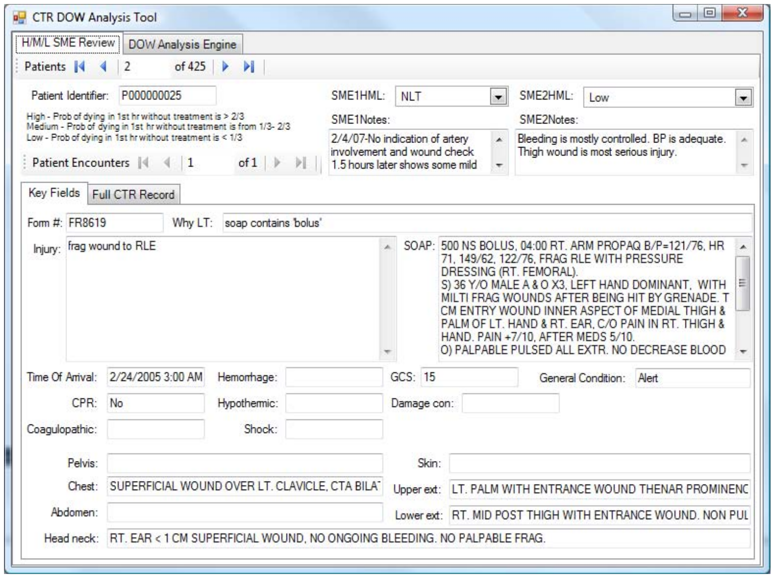


Figure 13. SME Review Tool.

5. All casualties that had at least one encounter, with adequate timing data and at least one encounter that indicated the patient may be LT, were passed on to the medical SMEs (Doug Lowe, Martin Hill) for independent evaluations. Each SME was asked to rate each patient as non-LT, H, M, L, or not enough information. After their initial reviews, they were asked to compare results and reconcile any differences.
6. This reconciled list of casualties and encounters was exported to an Excel file for a statistical analysis.

CTR EMED Data DOW PC Estimator Algorithm

Because we decided the Estimator Algorithm was not producing accurate enough results and we did not have time to continue improving it during this contract period, we stopped development on the Estimator Algorithm and concentrated on defining LT criteria instead (above) and having those results reviewed by SMEs. The algorithm below was not used to produce the current set of data; it is included only for completeness.

The following is an overview of the steps performed by the PC Estimator Algorithm in the order that they occur.

1. Catalog all words used in all analyzed text fields in the CTR EMED data. Any word that can be associated with a synonym is changed to that synonym (for instance, "pelvic" becomes "pelvis"). Some words are omitted (such as "and"), and some words are kept together (such as "lower" and anything that follows it, like "lower extremity"). Each word is given a weight based on the location in which it was found (for instance, words from the injury description are given higher weight than words from SOAP notes). See appendices for list of words, synonyms, weights, and other rules applied to the algorithm at this stage.
2. Parse out words for each PC description in a similar manner.

3. For each word that occurred in a CTR EMED record, see if that word is in the "instant PC match" list. If so, record the match with an extremely high weight to force selection of that PC later (see Appendix F for a list of instant-match words and their PCs)
4. For each word that occurred in a CTR EMED record, see if that word is in the "instant category match" list. If so, limit PC selections to that category. If not, look at all categories. Compare each word in the CTR EMED record with each word from each PC description. Record any matches with the weight assigned to the word in step 1.
5. For the same record, look at the list of words that are important to a PC (see Appendix F for list) and assign a NEGATIVE weight match for any word that did not appear in the CTR EMED record. For instance, if a PC description includes "brain" but that word was not in the CTR EMED record, a negative weight is assigned to try and push the Estimator away from that PC.
6. For each CTR EMED record, look at the anatomical categories chosen and choose a PC category based on the anatomical location if possible (see Appendix F for details of the criteria). If no anatomical location is recorded, do not limit the search to a particular category.
7. Assign an LT probability of "Definitely", "Probably", or "Maybe" to each CTR EMED record based on criteria in Appendix F.
8. For each CTR EMED record and each PC, divide the sum of the weights for every match in the CTR EMED record by the number of valid words in the PC description. This gives an index value that is theoretically more consistent between PCs that have different numbers of words in their description.
9. For each CTR EMED record, look at the index (computed in last step) for each match recorded in the PC category selected in step 6. If "Definitely" was selected for the CTR EMED record in step 7, limit the possible PCs further by selecting only from those that were classified as LT by SMEs.
10. Choose the PCs that have the top three indices as computed in step 8. In cases where there are multiple records for a single patient, choose the top three PCs among all those computed for EITHER record associated with the patient.

Appendix F. Parameters for Lexical Analysis

Any found word (in PC or CTR EMED record) is converted to the synonym word on the right of the equal sign. Word forms ending in "s" or "es" are automatically considered as well.

"rt" = "right"	"spinal" = "spine"
"lt" = "left"	"extrem", "extreme", "large", "massive", "through", "penetrating" = "severe"
"pelvic" = "pelvis"	"pneumohemothorax" = "lung"
"genitalia", "groin", "perineal", "scrotal" = "genital"	"nontraumatic" = "no trauma"
"extremity", "extremities" = "limb"	"superficial" = "slight"
"fracture", "break" = "broken"	"intracranial" = "scalp"
"amp", "amputate", "bka", "amputee" = "amputation"	"forearm" = "arm"
"cm" = "centimeter"	"wnd" = "wound"
"non", "not", "without" = "no"	"dislocation" = "dislocated"
"facial", "cheek", "cheeks" = "face"	"femur", "kneecap", "knee", "thigh", "patellae", "le" = "leg"
"meds", "medicine" = "medication"	"trunk", "flank", "torso", "disemboweled", "disembowel", "abdominal" = "abdomen"
"matter", "cerebral", "intercranial", "csf" = "brain"	"respiratory", "breath", "breathing" = "lung"
"contrib", "contributing", "contributes" = "contribute"	"neck", "skull", "temporal", "cranial", "cranium", "globe" = "head"
"eyelid", "eyeball", "intraocular" = "eye"	"pleural", "axilla", "thorax", "thoracic" = "chest"
"mandible" = "jaw"	

Omitted, Dependent, and Instant PC/Category Match Words

These words (and their synonyms, if applicable) are ignored during the matching process:

Assigned	With	But
And	Than	At
Or	The	Wound
To	All	From
Of	In	

These words are always kept before the word to their right (for instance, "no amputation" is kept together to keep "amputation" from matching in a PC description):

No	Lower	Upper

A CTR EMED record with one of these words is instantly matched to PC 17 (closest PC to "decapitation"). Note that decapitation was purposefully misspelled to capture misspellings in the CTR EMED data:

Decapitation	Decapitated	Decapetation

In cases where there is no anatomical record, the following words, if found, will force the estimator into the PC category specified. If more than one of these categories is found, it is forced into the MIW category:

head → Head	arm → Upper Limbs	pelvis → Abdomen & Pelvis
chest → Thorax	abdomen → Abdomen & Pelvis	spine → Spine
brain → Head	leg → Lower Limbs	

Key Indicator Words and Weights Assigned

These words are considered key indicators of an injury description, and are awarded extra weight as specified:

head +30	face, +30	eye, +30
abdomen, +30	chest, +30	amputation, +10

These fields in the CTR EMED record are assigned weights as well. Any word matched in these locations receives the weight specified:

MOI_1_8 (aka MOI1), +1	Burn_TBSA, +1	Injury_Category2, +100
soap, +1	abdomen, +50	pelvis, +50
MOI_9_16 (aka MOI2), +1	injury, +100	Injury_Category3, +100
head_neck, +50	upper_ext, +50	skin, +50
MOI_Burn, +1	Injury_Category1, +100	Injury_Category4, +100
chest, +50	lower_ext, +50	

The text representation of the anatomical locations (head, lower extremity, abdomen, etc.) selected also receives a weight bonus of 200.

There is also a list of words associated with PC descriptions that are considered key words describing the PC. If a CTR EMED record does NOT contain one of these words it is given a negative weight for that PC. This helps prevent matching of PCs that contain a single word that matched (like "perforating") but ignore a more important word (like "lung").

brain, -100	arm, -100	abdomen, -100
lung, -100	colon, -100	liver, -100
leg, -100	head, -100	spine, -100
chest, -100	spleen, -100	kidney, -100

Anatomical Location PC Estimators

The anatomical location, if recorded, is used to limit the PC Estimator Algorithm to a range of likely PCs. The anatomical location is a series of yes or no values in each of several categories. They are related to the PC category as shown below:

Anatomical record fields checked "Yes"

* Buttock (never selected)	* Thorax/Back or Chest
* Neck, Head, Face, Eye, or Ear	Lower Extremity
* Genitalia, Abdomen, or Pelvis	Upper Extremity
* Back	Two or more criteria with * (above)

Resulting PC category

Buttock (no such PC category)	Thorax
Head	Lower Limbs
Abdomen & Pelvis	Upper Limbs
Spine	Multiple Injury Wounds

Word and Word Combinations for Defining LT

(Established at NHRC TIM; refined 21 December 2006 by Doug Lowe)

Definite LT – Single Words or Acronyms

KIA	Packed Cells	Strider
DOW	Whole Blood	Striders
CPR	Factor 7	Striderous
Deceased	Hypovolemia	Airway Distress
Arrest	Hypovolemic	RSI
Died	Amputate	Bolus
Asystole	Amputation	Bolis
Asystolic	Intubate	WBB
PC	Intubated	FAST+
Packed Cells	ET	Unresponsive
RBC	ETT	Exsangunate
Red Blood	Vent	Exsanguination
Cells	Ventilator	Bagged

Definite LT – Strings or Word Combinations

Cardiopulmonary Resuscitation	Head Injury–Open	No or Neg Cardiac Activity
Walking Blood Bank	Open Head Injury	No Pulse
Pos FAST	Pupils Fixed and Dilated (or any combination all)	No Spontaneous (or Spont Breathing or Breaths)
Positive FAST	No BS	Flail Chest
Chest Tube	No or Neg Breath Sounds	Pinpoint Pupils

Definite LT – Data Fields with Text

Triage=Expectant	Complications=Cardiac Arrest or Death–Non-Preventable or	Date of Death=Any Entry

	Hypovolemia-Hemorrhage or Deceased or Aspiration Pneumonia or Pneumothorax or Coagulopathy	
Any Injury Category=Head Injury-Open (or any combination of all)	Coagulopat or Coagulopathic=Yes	Time of Death=Any Entry
General Condition=Unresponsive	Hemorrhage=II or III or IV,	Airway Management=ETT or Cric or Cricothyrotomy or Tracheostomy
CPR=Yes	Hyperthermic=Yes	Damage Control=Yes
Chest Wound Care=Chest Tube	Disp Type=KIA or DOW or Deceased	GSC≤10

Words in Chest

Sucking Chest Wound	Decreased Breath Sounds
Pheumothorax	Decreased or decr BS
Hemothorax	Chest Tube
Hemo/Pneumothorax	Chest Tubes
Hemo, Pneumo	Flail Chest

Head/Neck

Open Head Injury	Carbonaceous Airway
Second or Third Degree Burn	T&T (thru and thru)

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD MM YY)
15 May 09

2. REPORT TYPE
Technical Report

3. DATES COVERED (from – to)
June 2006–May 2009

4. TITLE AND SUBTITLE
Empirical Analysis of Operation Iraqi Freedom Combat Mortality Using the Navy-Marine Corps Combat Trauma Registry Expeditionary Medical Encounter Database for Applications to Tactical Medical Logistics Modeling and Simulation

5a. Contract Number: N/A
5b. Grant Number: N/A
5c. Program Element Number:
5d. Project Number:
5e. Task Number:
5f. Work Unit Number: 60315

6. AUTHORS
Mitchell, Ray, and Parker, Joe

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commanding Officer
Naval Health Research Center
140 Sylvester Rd
San Diego, CA 92106-3521

8. PERFORMING ORGANIZATION REPORT NUMBER
09-33

8. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

Commanding Officer
Naval Medical Research Center
503 Robert Grant Ave
Silver Spring, MD 20910-7500

Commander
Navy Medicine Support Command
P.O. Box 140
Jacksonville, FL 32212-0140

10. SPONSOR/MONITOR'S ACRONYM(S)
NMRC/NMSC

11. SPONSOR/MONITOR'S REPORT NUMBER(s)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Prepared in cooperation with Teledyne Brown Engineering and SAIC.

14. ABSTRACT

Reducing U.S. forces' battlefield mortality is a paramount concern. Methods to more accurately forecast medical needs, for example by tactical medical logistics modeling and simulation, are constantly being developed. The Naval Health Research Center's (NHRC) Tactical Medical Logistics (TML+) simulation tool continues to gain respect in the planning community. However, further testing and improvement is needed to confirm or strengthen its capabilities to deliver accurate and useful information to planners. The main objectives of this study were describe a data-mining research effort, and determine if a probability model could adequately describe the mortality events in a casualty's medical treatment flow as recorded in the NHRC Combat Trauma Registry (CTR) database. Confirming or updating expert medical doctor (MD) panel (2003) outcomes with a statistical analysis of empirical mortality results was an obvious research goal. Results indicate that using subject matter expert (SME) input to provide missing timing data is a legitimate way to strengthen biomedical science data records. Additionally, statistical analysis results showed a strong graphical agreement with results estimated by the 2003 MD SME panel, thus bolstering the efficacy of TML+ output.

15. SUBJECT TERMS
Weibull probability density function (PDF), Combat Trauma Registry, subject matter experts

16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	18a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Commanding Officer
UNCL	UNCL	UNCL	UU	40	**18b. TELEPHONE NUMBER (INCLUDING AREA CODE)** COMM/DSN: (619) 553-8429